УДК 519.111

И. О. Коряков

ЭССЕ О ДВУХ КОМБИНАТОРНЫХ ФУНКЦИЯХ, ПОДЧИНЕННЫХ ЧИСЛАМ КАТАЛАНА*

1. Преамбула

В задачах перечисления часто встречается следующая схема. Имеется последовательность $(M_n)_{n\in\mathbb{N}}$ множеств комбинаторных объектов – модель для числовой последовательности $(G_n)_{n\in\mathbb{N}}$, т. е. $|M_n| = G_n$ для всех n, а каждое M_n является дизъюнктным объединением подмножеств M_{nk} ($1 \leqslant k \leqslant n$ или $0 \leqslant k \leqslant n$), причем $|M_{nk}| = g(n, k)$. Тогда мы называем функцию $g(n, k)$ *подчиненной* функции G_n, в частности $G_n = \sum_k g(n, k)$. Например, $G_n = 2^n$ и $g(n, k) = \binom{n}{k}$; G_n – числа Белла и $g(n, k)$ – числа Стирлинга 2-го рода; $G_n = n!$ и $g(n, k)$ – числа Стирлинга 1-го рода без знака; $G_n = n!$ и $g(n, k) = A(n, k)$ – числа Эйлера и т. п.

Числа Каталана $C_n = \frac{1}{n+1}\binom{2n}{n}$ особенно богаты разнообразными моделями, и в этой статье на основе ряда элементарных моделей чисел C_n обсуждаются две подчиненные им комбинаторные функции: (безымянные) числа $C(n, k)$ и числа Рюньона $R(n, k)$. Я старался указать источники результатов, но это не всегда удавалось, так что в итоге я не знаю, какие результаты являются новыми. Но для целей статьи это не слишком важно. Все приводимые в работе утверждения доказываются с помощью простейших биекций; второй план статьи – тривиализация доказательств и демонстрация преимуществ биективного подхода. Еще один мотив – желание уточнить некоторые комбинаторные определения; в частности, определения сканирования деревьев. Основные модели – слова и деревья.

2. Словарное введение

Всюду ниже $\mathbb{N} = \{0, 1, 2, \ldots\}$, $\langle n\rangle = \{1, 2, \ldots, n\}$ для $n \in \mathbb{N}$, $\mathbb{Z}$ – аддитивная группа целых чисел. Необходимые сведения о формальных языках и кодах можно найти в [1, гл. 5]. Если u – префикс слова w, то мы пишем $u \leqslant w$. Буква e обозначает пустое слово.

*Работа выполнена при поддержке программы «Развитие научного потенциала высшей школы», проект № 2.1.1/3537.

© Коряков И. О., 2010

Зафиксируем алфавиты $X = \{x, \overline{x}\}$ и $Y = \{x, \overline{x}, \ell, r\}$ и рассмотрим морфизм $\delta : Y^* \to \mathbb{Z}$, заданный равенствами $\delta(x) = 1$, $\delta(\overline{x}) = -1$, $\delta(\ell) = \delta(r) = 0$. Множество $\mathcal{M}$ всех слов $w \in Y^*$, удовлетворяющих условиям

$$\delta(w) = 0 \quad \text{и} \quad u \leqslant w \Rightarrow \delta(u) \geqslant 0,$$

называется (двухцветным) *языком Моцкина*, а $\mathcal{D} = \mathcal{M} \cap X^*$ – (ограниченным) *языком Дика*. Очевидно, слова Дика имеют четную длину и получаются из слов Моцкина стиранием букв ℓ и r. Положим $\mathcal{M}_n = \mathcal{M} \cap Y^n$, $\mathcal{D}_n = \mathcal{D} \cap X^{2n}$. Слово Дика *неразложимо*, если его нельзя представить в виде конкатенации других слов Дика. Неразложимые слова Дика образуют бипрефиксный код $x\mathcal{D}\overline{x}$. Нетрудно доказать следующие три леммы.

Лемма 1. *Для любого непустого $w \in \mathcal{D}$ существуют такие однозначно определенные $u, v \in \mathcal{D}$, что $w = xu\overline{x}v$. Обратно, любое слово указанного вида принадлежит $\mathcal{D}$.*

Из леммы 1 вытекает

Лемма 2. *Для любого $w \in \mathcal{D}$ существует единственное $k \in \mathbb{N}$ такое, что $w \in (x\mathcal{D}\overline{x})^k$.*

Лемма 3. *Для любого непустого $w \in \mathcal{M}$ существуют такие однозначно определенные $u, v \in \mathcal{M}$, что либо $w = lu$, либо $w = rv$, либо $w = xu\overline{x}v$. Обратно, любое слово указанного вида принадлежит $\mathcal{M}$.*

Пусть A – конечный или счетный алфавит, а морфизм $\delta : A^* \to \mathbb{Z}$ таков, что $\delta \geqslant -1$ на A, причем $\delta^{-1}(-1) \cap A \neq \varnothing$. Множество L всех слов $w \in A^*$ таких, что $\delta(w) = -1 \wedge (u \leqslant w \Rightarrow \delta(u) \geqslant 0)$, называется *сильно префиксным кодом относительно* δ или просто δ-*кодом* (см. [2]). Примерами δ-кодов служат языки $\mathcal{D}\overline{x}$ и $\mathcal{M}\overline{x}$. Еще один пример – *язык Лукасевича* $\mathcal{L}$ над $Z = \{x_n \mid n \in \mathbb{N}\}$, определяемый как δ-код при $\delta(x_n) = n - 1$, $n \in \mathbb{N}$. Положим $\mathcal{L}_n = \mathcal{L} \cap Z^n$. Известна

Лемма 4. *Пусть $L \subseteq Z^*$; $L = \mathcal{L}$ тогда и только тогда, когда*

$$L = \bigcup_{n \in \mathbb{N}} x_n L^n.$$

Слова u, v называются *сопряженными* (пишем $u \sim v$), если одно получается из другого циклической перестановкой букв. Язык $Q \subseteq A^*$ *устойчив* относительно сопряжения, если $(u \in Q \wedge v \sim u) \Rightarrow v \in Q$. Положим $Q_n = Q \cap A^n$.

Обобщенная формула Рэйни (ОФР). *Если L – δ-код над алфавитом A и язык $Q \subseteq A^*$ устойчив относительно сопряжения, то при $k, n > 0$*

$$\left|L^k \cap Q_n\right| = \frac{k}{n}\left|\delta^{-1}(-k) \cap Q_n\right|.$$

Эта формула получена в [3] для $L = \mathcal{L}$ и $Q = A^*$, а в полном объеме – в [2].

3. Древесное введение

В этой статье *все деревья – конечные корневые*; используются обычные для узлов корневого дерева термины: *сын, отец, лист, брат, потомок* (понятия «сын», «отец», «левый сын» и т. п. фигурируют в двух ипостасях – как бинарные предикаты и как унарные; читатель разберется, что имеется в виду в каждом случае). Всякое поддерево отождествляется со множеством его узлов; в частности, мощность $n = |T|$ дерева T – это число его узлов, тогда T называется *n-деревом*. Если a – узел дерева T, то $T(a)$ – поддерево, состоящее из a и всех его потомков. Мы рассматриваем деревья двух сортов – бинарные и упорядоченные.

Дерево называется *бинарным*, если каждый сын либо левый, либо правый, причем каждый узел имеет не более одного левого и не более одного правого сына. Обозначим через $\mathcal{B}$ множество всех бинарных деревьев, а через $\mathcal{B}_n$ – его подмножество n-деревьев, $n \in \mathbb{N}$.

Пусть $T \in \mathcal{B}$ непусто; положим $LT = T(a)$, если a – левый сын корня, и $LT = \varnothing$, если у корня нет левого сына; двойственно определяется поддерево RT. Мы пишем $T \simeq (LT, RT)$. Таким образом, для любого $T \in \mathcal{B}$ либо $T = \varnothing$, либо $T \simeq (L, R)$, где $L, R \in \mathcal{B}$; это *рекурсивное представление* дерева T.

Дерево называется *упорядоченным*, если оно непусто и для каждого узла a множество $S(a)$ его сыновей линейно упорядочено. Этот порядок часто называют *старшинством*, записывая (и рисуя) братьев слева направо по убыванию старшинства. Обозначим через $\mathcal{T}$ множество всех упорядоченных деревьев, а через $\mathcal{T}_n$ – его подмножество n-деревьев, $n > 0$.

Пусть $(a_1, \ldots, a_k)$ – список сыновей корня дерева $T \in \mathcal{T}$. Если этот список пуст, пишем $T \simeq (\varnothing)$; иначе, положив $T_i = T(a_i)$, $i \in \langle k \rangle$, пишем $T \simeq (T_1, \ldots, T_k)$, – справа записан упорядоченный лес, это стандартное *рекурсивное представление* дерева T. Введем *новое рекурсивное представление*, при котором не надо заботиться о количестве сыновей. Если $T = \varnothing$, положим $LT = \varnothing$, $RT = T$; если же $T \simeq (T_1, \ldots, T_k)$, $k > 0$, положим $LT = T_1$, $RT \simeq (T_2, \ldots, T_k)$ и пишем $T \cong (LT, RT)$. Таким образом, для любого $T \in \mathcal{T}$ имеем $T \cong (LT, RT)$, где RT – упорядоченное дерево, полученное из T «отрыванием» LT.

Обход дерева «слева–направо» указывает общее направление: левое (старшее) поддерево обрабатывается раньше правого (младшего). *Сканирование* (обычно чтение меток, но иногда их присваивание и т. п.) требует специального алгоритма. Мы различаем три режима узлового сканирования, которые могут реализовываться одновременно: «сверху вниз» (act1), «симметрично» (act2) и «снизу вверх» (act3). Традиционный «обход» (англ. *transversal* или *transversing*, см., например, [4, п. 2.3.1]) не всегда допускает одновременность, требуя посещения корня лишь единожды. Приведем рекурсивный алгоритм NSCANB узлового (N) сканирования (чтения меток) бинарных (B) деревьев. Предполагается, что каждый узел может быть снабжен одной, двумя или тремя метками; рассмотрим тройную метку (ℓ, c, r); результат – слово w.

NSCANB:

$w := e$;
SCAN(T) {if $T = \varnothing$ then exit;
 act1: $w := w\ell$; SCAN(LT);
 act2: $w := wc$; SCAN(RT);
 act3: $w := wr$}.

Аналогично строится алгоритм NSCANO узлового сканирования упорядоченных деревьев. Пример будет дан в предложении 3.

Не лучше обстоит дело с реберным сканированием. Например, в [5] применяется «обход (walk around) [плоского] дерева, начиная с корня, против часовой стрелки»; в [6, гл. 20] для этого используется «червяк» (хотя на рисунке изображена многоножка). Укажем алгоритм ESCANO реберного (E) сканирования упорядоченных (O) деревьев в режимах «сверху вниз» (act1) и «симметрично» (act2), основанный на новом рекурсивном представлении:

ESCANO:

$w := e$;
SCAN(T) {if $LT = \varnothing$ then exit;
 act1: $w := wl$; SCAN(LT);
 act2: $w := wc$; SCAN(RT)}.

Аналогично строится алгоритм ESCANB реберного сканирования бинарных деревьев. Пример будет дан в предложении 8.

4. Числа $C(n,k) = \dfrac{k}{2n-k}\dbinom{2n-k}{n}$

Предложение 1. (1) *Число слов из $\mathcal{D}_n$, являющихся произведением k неразложимых слов Дика, равно $C(n,k)$.*
(2) $|\mathcal{D}_n| = C_n$.

Доказательство. Перечисляемым множеством является

$$\mathcal{D}_n \cap (x\mathcal{D}\overline{x})^k = (x\mathcal{D}\overline{x})^k \cap X^{2n}.$$

Очевидна биекция $x\mathcal{D}\overline{x}$ на δ-код $\mathcal{D}\overline{x}$. В силу этой биекции и ОФР

$$\left|(x\mathcal{D}\overline{x})^k \cap X^{2n}\right| = \left|(\mathcal{D}\overline{x})^k \cap X^{2n-k}\right| = \frac{k}{2n-k}\left|\delta^{-1}(-k) \cap X^{2n-k}\right|.$$

Чтобы сформировать слово из $\delta^{-1}(-k)\cap X^{2n-k}$, нужно из $2n-k$ мест выбрать n мест для $\overline{x}$; поэтому $\left|\delta^{-1}(-k) \cap X^{2n-k}\right| = \binom{2n-k}{n}$.

(2) следует из (1):

$$\left|\mathcal{D}_n\right| = \left|x\mathcal{D}_n\overline{x}\right| = \left|x\mathcal{D}\overline{x} \cap X^{2n+2}\right| = \frac{1}{2n+1}\binom{2n+1}{n+1} = C_n.$$

Утверждение (2) хорошо известно. Утверждение (1) можно вывести из одной теоремы в [5], гласящей, что число p-списков слов Дика суммарной длины $2n$ равно $\big(p/(2n+p)\big)\dfrac{(2n+p)!}{n!(n+p)!}$ $[= C(n+p,p)]$.

Предложение 2. (1) $|\mathcal{T}_{n+1}| = C_n$.
(2) *Число деревьев из $\mathcal{T}_{n+1}$, корень которых имеет k сыновей, равно $C(n,k)$.*

Доказательство. (1) Присвоим каждому ребру дерева $T \in \mathcal{T}$ двойную метку $(x, \overline{x})$ и применим алгоритм ESCANO, одновременно сканируя метку x сверху вниз, а метку $\overline{x}$ симметрично. Пусть $\varphi(T)$ – результат сканирования. Если $LT = \varnothing$, то $\varphi(T) = e$; если $LT \neq \varnothing$ и $T \cong (LT, RT)$, то $\varphi(T) = x\varphi(LT)\overline{x}\varphi(RT)$. Индукция с использованием леммы 1 доказывает, что φ – биекция $\mathcal{T}$ на $\mathcal{D}$, причем $\varphi(\mathcal{T}_{n+1}) = \mathcal{D}_n$. Поэтому (1) вытекает из предложения 1(2).

(2) Из предыдущего следует, что если $T \simeq (T_1, \ldots, T_k)$, то

$$\varphi(T) = (x\varphi(T_1)\overline{x}) \ldots (x\varphi(T_k)\overline{x}) \in (x\mathcal{D}\overline{x})^k.$$

Лемма 2 показывает, что корень дерева T имеет k сыновей тогда и только тогда, когда $\varphi(T) \in (x\mathcal{D}\overline{x})^k$. Теперь (2) следует из предложения 1(1).

Утверждение, равносильное (1), приписывают Кэли (см. [4] или [6]). В [7, с. 129] с помощью формулы Лагранжа–Бюрмана доказывается утверждение, равносильное (2): «число плоских деревьев с висячим корнем, с n некорневыми вершинами, у которых вершина, смежная с корнем, имеет степень m, равно $\frac{m-1}{n-1}\binom{2n-m-2}{n-2}$», что, в свою очередь, совпадает с $C(n-1, m-1)$.

Предложение 3. (1) $|\mathcal{L}_{n+1}| = C_n$.
(2) *Число слов из* $\mathcal{L}_{n+1}$, *начинающихся с* k, *равно* $C(n,k)$.

Доказательство. (1) Узел дерева $T \in \mathcal{T}$, имеющий k сыновей, пометим буквой $x_k \in Z$. Пусть $\lambda(T)$ – результат узлового сканирования сверху вниз дерева T. Тогда если $T \simeq (\varnothing)$, то корень не имеет сыновей и $\lambda(T) = x_0$; если $T \simeq (T_1, \ldots, T_k)$, $k \geqslant 1$, то $\lambda(T) = x_k \lambda(T_1) \ldots \lambda(T_k)$. Индукция с использованием леммы 4 доказывает, что λ – биекция $\mathcal{T}$ на $\mathcal{L}$, причем $\lambda(\mathcal{T}_{n+1}) = \mathcal{L}_{n+1}$.

(2) Очевидно, словам из перечисляемого множества отвечают деревья с k сыновьями корня. Поэтому все утверждение вытекает из предложения 2.

Путь длины n в плоской целочисленной решетке $\mathbb{Z} \times \mathbb{Z}$ – это список $\pi = (P_0, P_1, \ldots, P_n)$ точек решетки; векторы $\pi_i = P_i - P_{i-1}$, $i \in \langle n \rangle$, – *шаги* пути π. Если точка P_0 фиксирована (обычно $P_0 = (0,0)$), то π однозначно определен своими шагами, и мы пишем $\pi = (\pi_1, \ldots, \pi_n)$. Назовем *стандартным* путь с шагами $(1,0)$ («вправо») и $(0,1)$ («вверх»). *Диагональ* решетки – это прямая $y = x$. Через $\mathcal{P}$ (соответственно $\mathcal{P}_n$) обозначим множество всех стандартных путей, начинающихся в точке $(0,0)$, лежащих не ниже диагонали и заканчивающихся на ней (соответственно в точке (n,n)).

Предложение 4. (1) $|\mathcal{P}_n| = C_n$.
(2) *Число путей из* $\mathcal{P}_n$ *с* k *точками на диагонали равно* $C(n,k)$.

Доказательство. (1) Соответствие $(1,0) \mapsto \overline{x}$, $(0,1) \mapsto x$ задает биекцию $\mathcal{P} \to \mathcal{D}$, отображающую $\mathcal{P}_n$ на $\mathcal{D}_n$.

(2) вытекает из предложения 1, так как каждому участку пути между соседними точками, лежащими на диагонали, отвечает неразложимое слово Дика.

Утверждение (1) можно вывести из теоремы Феллера о баллотировке (см. [8], где фактически рассматриваются «диагональные» пути с шагами $(1,1)$ и $(1,-1)$).

Преобразование σ множества $\langle n \rangle$ называется *экстенсивным*, если $\sigma(i) \geqslant i$ для всех $i \in \langle n \rangle$, и *возрастающим*, если $i < j \Rightarrow \sigma(i) \leqslant \sigma(j)$. Как обычно, число $i \in \langle n \rangle$ – *неподвижная точка* преобразования σ, если $\sigma(i) = i$. Множество $EI(n)$ всех экстенсивных возрастающих преобразований на $\langle n \rangle$ образует моноид относительно композиции.

Предложение 5. (1) $|EI(n)| = C_n$.
(2) *Число преобразований из* $EI(n)$ *с* k *неподвижными точками равно* $C(n,k)$.

Доказательство. (1) Определим отображение из $\mathcal{P}_n$ в $EI(n)$, $\pi \mapsto \sigma$, положив $\sigma(i)$ равным наименьшей ординате точек пути π с абсциссой i, $i \in \langle n \rangle$. Обратное отображение: для заданного σ введем расширенный график

$$\Gamma(\sigma) = \{(0,0)\} \cup \{(i, \sigma(i)) \mid i \in \langle n \rangle\};$$

тогда π – это $\Gamma(\sigma)$ с добавленными для каждого i, $0 \leqslant i < n$, такого, что $y = \sigma(i) < \sigma(i+1) = z$, точками $(i, y+1), \ldots, (i, z)$.

(2) Неподвижные точки преобразования σ отвечают точкам пути π, лежащим на диагонали и отличным от $(0,0)$. Поэтому утверждение следует из предложения 4.

Левой ветвью бинарного дерева назовем наименьшее поддерево, содержащее корень и удовлетворяющее условию: вместе с каждым узлом оно содержит левого сына этого узла. Двойственно определяется *правая ветвь*.

Предложение 6. (1) $|\mathcal{B}_n| = C_n$.
(2) *Число деревьев из $\mathcal{B}_n$ с правой ветвью, содержащей k узлов, равно $C(n,k)$.*

Доказательство. (1) Рассмотрим алгоритм [4, п. 2.3.2], перестраивающий бинарное n-дерево в упорядоченное $(n+1)$-дерево. Он переоформляет «семейные» отношения: a) каждый левый сын объявляется старшим сыном; b) каждый правый сын объявляется младшим братом своего (прежнего) отца; c) к полученному упорядоченному лесу добавляется новый корень. Это биекция $\mathcal{B}_n$ на $\mathcal{T}_{n+1}$.

(2) Узлам правой ветви исходного дерева отвечают сыновья нового корня. Поэтому предложение 6 следует из предложения 2.

Конечно, в предложении 6 правую ветвь можно заменить левой.

Предложение 7. *Число слов из $\mathcal{D}_n$ с максимальным префиксом из $\{x\}^*$ длины k равно $C(n,k)$.*

Доказательство. Пометим каждый узел дерева $T \in \mathcal{B}$ двойной меткой $(x, \overline{x})$. Пусть $\psi(T)$ – результат одновременного сканирования метки x сверху вниз и метки $\overline{x}$ симметрично. Алгоритм NSCANB дает $\psi(\varnothing) = e$; если $T \simeq (LT, RT)$, то $\psi(T) = x\psi(LT)\overline{x}\psi(RT)$. Индукция с использованием леммы 1 доказывает, что ψ – биекция $\mathcal{B}_n$ на $\mathcal{D}_n$ для любого $n \in \mathbb{N}$. Буквы x до первого вхождения $\overline{x}$ в $\psi(T)$ получаются в точности при сканировании левой ветви дерева T. Поэтому утверждение вытекает из предложения 6.

Заметим, что можно сканировать x симметрично, $\overline{x}$ снизу вверх; этим определяется другая биекция $\mathcal{B}_n$ на $\mathcal{D}_n$, позволяющая заменить «префикс» на «суффикс», а x на $\overline{x}$.

Предложение 8. (1) $|\mathcal{M}_{n-1}| = C_n$ *для* $n \geqslant 1$.
(2) *Число слов из* $\mathcal{M}_{n-1}$ *с максимальным префиксом из* $\{\ell, x\}^*$ *длины* k *равно* $C(n,k)$.

Доказательство. (1) Пометим ребра непустого дерева $T \in \mathcal{B}$ буквами алфавита Y: буквой ℓ (соответственно r) – ребро, ведущее к левому (соответственно правому) сыну, не имеющему брата; буквами x и $\overline{x}$ – ребра, ведущие соответственно к левому и правому сыновьям, когда имеются оба брата. Пусть $\mu(T)$ – результат реберного сканирования дерева T сверху вниз. Тогда если $T \simeq (\varnothing, \varnothing)$, то $\mu(T) = e$; если $|T| \geqslant 2$ и $T \simeq (LT, RT)$, то

$$\mu(T) = \begin{cases} \ell\mu(LT), & \text{если} \quad RT = \varnothing; \\ r\mu(RT), & \text{если} \quad LT = \varnothing; \\ x\mu(LT)\overline{x}\mu(RT), & \text{если} \quad LT, RT \neq \varnothing. \end{cases}$$

Индукция с использованием леммы 3 доказывает, что μ – биекция $\mathcal{B}_n$ на $\mathcal{M}_{n-1}$ для любого $n \geqslant 1$.

(2) Буквы x и ℓ до первых вхождений $\overline{x}$ или r в $\mu(T)$ появляются в точности при сканировании левой ветви дерева T.

Заметим, что если бинарное дерево полно (т. е. непусто и каждый отец имеет двух сыновей), то μ определяет биекцию $\mathcal{CB}_n \to \mathcal{D}_n$, где $\mathcal{CB}_n$ – множество всех полных бинарных $(2n+1)$-деревьев. Она совпадает с композицией двух биекций – «дефолиации» (т. е. обрывания всех листьев полного дерева) и биекции ψ из доказательства предложения 7.

Обозначим через $\mathcal{B}_n(k)$ множество деревьев из $\mathcal{B}_n$ с k узлами в правой ветви.

Предложение 9. *Число* k*-списков бинарных деревьев с суммарным числом* m *узлов равно* $C(m+k,k)$.

Доказательство. Пусть $(a_1, \ldots, a_k)$ – список узлов правой ветви дерева $T \in \mathcal{B}_{m+k}(k)$, где a_{i+1} – правый сын узла a_i; $T_i = LT(a_i)$, $i \in \langle k \rangle$. Тогда $\sum_{i \in \langle k \rangle} |T_i| = m$ и соответствие $T \mapsto (T_1, \ldots, T_k)$ есть биекция $\mathcal{B}_{m+k}(k)$ на перечисляемое множество. Нужный результат вытекает из предложения 6.

Обратимся к внутренним связям чисел $C(n,k)$. Заметим сначала, что $C(n,1) = C(n,2) = C_{n-1}$, и положим $C(n,0) = 0$.

Предложение 10. *Для* $n \geqslant 1$ *и* $k \in \langle n \rangle$

$$C(n+1,k+1) = C(n+1,k) - C(n,k-1).$$

Доказательство. При $k = 1$ равенство верно. Пусть $k > 1$. Определим отображение

$$\mathcal{B}_{n+1}(k) \to \mathcal{B}_n(k-1) \cup \mathcal{B}_{n+1}(k+1), \quad T \mapsto T',$$

следующим образом. Пусть a – нижний узел правой ветви дерева T. Если a – лист, оборвем его, получая $\mathcal{B}_n(k-1)$. Если a не лист, то пусть b – его (левый) сын. Построим дерево $T' \in \mathcal{B}_{n+1}(k+1)$ перестройкой в $T(a)$: объявим b правым сыном узла a в T', оторвем от $T(b)$ поддерево $RT(b)$ и приклеим его к a слева, т. е. положим $LT'(a) = RT(b)$. Манипуляции, восстанавливающие status quo, очевидны; поэтому введенное отображение – биекция.

Имея очевидное равенство $C_n = \sum_{k=1}^{n} C(n,k)$, получим обратное преобразование. Положим $C(0,k) = 0$.

Предложение 11. *Для любых* $n \geqslant 1$, $k \geqslant 0$ *выполняется равенство*

$$C(n+1,k+1) = \sum_{m=0}^{\lfloor k/2 \rfloor} (-1)^m \binom{k-m}{m} C_{n-m}.$$

Доказательство. При $k = 0$ обе части равенства дают C_n. При $k > 0$, в силу предложения 10 и индуктивной гипотезы,

$$\begin{aligned} C(n+1,k+1) &= C(n+1,k) - C(n,k-1) = \\ &= \sum_{m \geqslant 0} (-1)^m \binom{k-1-m}{m} C_{n-m} - \sum_{\ell \geqslant 0} (-1)^\ell \binom{k-2-\ell}{\ell} C_{n-1-\ell}. \end{aligned}$$

Заменяя ℓ на $m-1$, получаем

$$\begin{aligned} C(n+1,&k+1) = \\ &= \sum_{m \geqslant 0} (-1)^m \binom{k-1-m}{m} C_{n-m} - \sum_{m \geqslant 1} (-1)^{m-1} \binom{k-1-m}{m-1} C_{n-m} = \\ &= C_n + \sum_{m \geqslant 1} (-1)^m \left[\binom{k-1-m}{m} + \binom{k-1-m}{m-1} \right] C_{n-m} = \\ &= \sum_{m \geqslant 0} (-1)^m \binom{k-m}{m} C_{n-m}. \end{aligned}$$

Наконец, из предложения 10 следует, что

$$C_{n-1} = C(n,1) = C(n,2) > C(n,3) > \ldots > C(n,n) = 1.$$

5. Числа Рюньона $R(n,k)$

Эти числа в виде r_{nm} ($= R(n,m+1)$ в нашем обозначении) появляются в [9, п. 1.4] без конкретных ссылок. Для $n \geqslant 1$ и $k \in \langle n \rangle$

$$R(n,k) = \frac{1}{n}\binom{n}{k}\binom{n}{k-1} = \frac{1}{k}\binom{n}{k-1}\binom{n-1}{k-1} = \frac{1}{n-k+1}\binom{n}{k}\binom{n-1}{k-1}.$$

Таблица чисел $R(n,k)$ похожа на треугольник Паскаля: $R(n,n) = R(n,1) = 1$, $R(n,n-1) = R(n,2) = \binom{n}{2}$; имеется симметрия: $R(n,n-k+1) = R(n,k)$; выполняются условия унимодальности:

$$R(n,1) < \ldots < R\left(n, \left\lfloor \frac{n+1}{2} \right\rfloor\right) = R\left(n, \left\lfloor \frac{n+2}{2} \right\rfloor\right) > \ldots > R(n,n).$$

Все это легко проверить непосредственно. Положим также $R(n,0) = 0$.

Полноты ради воспроизведем (из [2]) доказательство следующего утверждения.

Предложение 12. *Для $n \geqslant 1$ и $k \in \langle n \rangle$ число всех слов из $\mathcal{M}_{n-1}$ с суммарным числом $k-1$ вхождений букв x и ℓ равно $R(n,k)$.*

Доказательство. Пусть $d_a(w)$ обозначает число вхождений буквы a в слово w, а язык $Q \subseteq Y^*$ определен условием: $w \in Q$, если и только если

$$d_x(w) + +d_\ell(w) = k-1 \quad \text{и} \quad d_{\overline{x}}(w) = d_x(w) + 1.$$

Перечисляемое множество равномощно множеству $\mathcal{M}\overline{x} \cap Q_n$, где $\mathcal{M}\overline{x}$ – δ-код. Чтобы сформировать слово из $\delta^{-1}(-1) \cap Q_n$, нужно из n мест выбрать $k-1$ мест для размещения x и ℓ; из этих $k-1$ мест выбрать $k-1-i$ мест для ℓ (или i мест для x); из оставшихся $n-k+1$ мест выбрать $i+1$ мест для $\overline{x}$; наконец, просуммировать по i. ОФР и тождество Вандермонда завершают доказательство:

$$\left|\mathcal{M}\overline{x} \cap Q_n\right| = \frac{1}{n}\left|\delta^{-1}(-1) \cap Q_n\right| =$$
$$= \frac{1}{n}\binom{n}{k-1}\sum_{i=0}^{k-1}\binom{k-1}{k-1-i}\binom{n-k-1}{i+1} = \frac{1}{n}\binom{n}{k-1}\binom{n}{k}.$$

Разумеется, в этом предложении буквы x и $\overline{x}$, а также ℓ и r взаимозаменяемы.

Предложение 13. *Для $n \geqslant 1$ и $k \in \langle n \rangle$ число деревьев из $\mathcal{B}_n$, имеющих $k-1$ левых [правых] сыновей, равно $R(n,k)$.*

Доказательство. Воспользуемся биекцией $\mu : \mathcal{B}_n \to \mathcal{M}_{n-1}$ из доказательства предложения 8. Левым сыновьям отвечают буквы x и ℓ, поэтому результат следует из предложения 12.

Делитель $x\overline{x}$ слова Дика называется *пиком.*

Предложение 14. *Для $n \geqslant 1$ число слов из $\mathcal{D}_n$, имеющих k пиков, равно $R(n,k)$.*

Доказательство. Рассмотрим биекцию $\psi : \mathcal{B} \to \mathcal{D}$ из доказательства предложения 7 и для $T \in \mathcal{B}$ обозначим через $\pi(T)$ число пиков в слове $\psi(T)$, а через $\rho(T)$ – число правых сыновей в T. Покажем индуктивно, что для непустого T число пиков $\pi(T) = \rho(T) + 1$; тогда требуемый результат будет следовать из предложения 13.

Если $|T| = 1$, то $\rho(T) = 0$, $\psi(T) = x\overline{x}$, $\pi(T) = 1 = \rho(T) + 1$. Пусть $|T| \geqslant 2$ и $T \simeq (LT, RT)$. Рассмотрим случай, когда $LT, RT \neq \varnothing$. Тогда

$$\rho(T) = \rho(LT) + \rho(RT) + 1, \qquad \psi(T) = x\psi(LT)\overline{x}\psi(RT)$$

и явно выписанные вхождения x и $\overline{x}$ не входят в пики, т. е.

$$\pi(T) = \pi(LT) + \pi(RT) = (\rho(LT) + 1) + (\rho(RT) + 1) = \rho(T) + 1.$$

Случаи $LT = \varnothing$ или $RT = \varnothing$ рассматриваются аналогично.

Полученный результат доказывался (довольно сложно) многими математиками. Мне известны имена G. Kreweras (1970), D. Goujou-Beauchamps (1975), T. V. Narayana (1979).

Путь из множества $\mathcal{P}$ (см. предложение 4) наглядно представляет собой лестницу, каждая *ступень* которой есть пара шагов («вверх», «вправо»). Предложение 14 вместе с биекцией $\mathcal{P} \to \mathcal{D}$ из доказательства предложения 4 влекут за собой

Предложение 15. *Число путей из $\mathcal{P}_n$ с k ступенями равно $R(n,k)$.*

Число $i \in \langle n \rangle$ – *точка роста* преобразования $\sigma \in EI(n)$ (см. предложение 5), если $\sigma(i) > \sigma(i+1)$ (полагая $\sigma(0) = 0$, считаем 1 точкой роста). *Ранг* $\mathrm{rg}(\sigma)$ – это мощность множества значений σ на $\langle n \rangle$.

Предложение 16. *Число преобразований из $EI(n)$, имеющих k точек роста (или, что то же самое, имеющих ранг k), равно $R(n,k)$.*

Доказательство. Относительно биекции $\mathcal{P}_n \to EI(n)$, $\pi \mapsto \sigma$, из доказательства предложения 5 точки роста σ наглядно отвечают ступеням «лестницы» π. Но можно воспользоваться биекцией $EI(n)$ на $\mathcal{D}_n$:

$$\sigma \mapsto x^{\sigma(1)}\overline{x}x^{\sigma(2)-\sigma(1)}\overline{x}\dots x^{\sigma(n)-\sigma(n-1)}\overline{x};$$

тогда точки роста σ отвечают пикам соответствующего слова Дика, и результат вытекает из предложения 14.

Очевидное равенство $\sigma(\langle n\rangle) = \{\sigma(i) \mid i$ – точка роста$\}$ обусловливает переход к $\mathrm{rg}(\sigma)$.

Предложение 17. *Число деревьев из $\mathcal{T}_{n+1}$ с k листьями [отцами] равно $R(n,k)$.*

Доказательство. Воспользуемся биекцией $\varphi : \mathcal{T}_{n+1} \to \mathcal{D}_n$ из доказательства предложения 2 и первым рекурсивным представлением упорядоченных деревьев. Пусть $T \in \mathcal{T}_{n+1}$. Докажем по индукции, что число листьев в T равно числу пиков в $\varphi(T)$; тогда результат будет следовать из предложения 14.

Если $n = 0$, то $\varphi(T) = e$ (в T нет листьев, в $\varphi(T)$ нет пиков). Если $n > 0$, то $T \simeq (T_1,\dots,T_m)$ для некоторых $m \geqslant 1$ и $T_i \in \mathcal{T}$; тогда $\varphi(T) = (x\varphi(T_1)\overline{x})\dots(x\varphi(T_m)\overline{x})$. Очевидно, все листья в T – это листья поддеревьев $T_1,\dots,T_m$, а все пики в $\varphi(T)$ – это пики слов $\varphi(T_1),\dots,\varphi(T_m)$.

Утверждение об отцах следует из симметрии чисел Рюньона.

Доказанное предложение вместе с биекцией из доказательства предложения 3 влекут

Предложение 18. *Число слов из $\mathcal{L}_{n+1}$ с k вхождениями буквы x_0 равно $R(n,k)$.*

Ранжируем бинарное дерево, определяя (левый) *ранг* r узлов: ранг корня равен 1; если a – левый сын узла b, то $r(a) = r(b)+1$, а если правый, то $r(a) = r(b)$.

Предложение 19. *При $n > 0$ число деревьев из $\mathcal{B}_n$ с k узлами нечетного ранга равно $R(n,k)$.*

Доказательство. Рассмотрим также ранжирование, определенное тем же способом с единственным изменением: ранг корня равен 0. Пусть $\mathcal{B}^i$ обозначает множество бинарных деревьев с рангом корня $i \in \{0, 1\}$. Введем псевдоранг $pr(a) = r(a) \bmod 2$ и инволютивное отображение $\pi : \mathcal{B}^i \to \mathcal{B}^{1-i}$, меняющее псевдоранг i на $1-i$. Для $T \in \mathcal{B}^0 \cup \mathcal{B}^1$ обозначим через $\rho(T)$ число узлов в T псевдоранга 1, а через $\lambda(T)$ – число (не зависящее от псевдоранга) левых сыновей в T.

Каждое непустое дерево $T \in \mathcal{B}^i$ имеет очевидное рекурсивное представление $T \simeq (L, i, R)$, где i – псевдоранг корня и, значит, $L \in \mathcal{B}^{1-i}$, $R \in \mathcal{B}^i$. Определим рекурсивно отображение $* : \mathcal{B}^i \to \mathcal{B}^i$, $i = 0, 1$, положив $\varnothing^* = \varnothing$ и

$$(L, 1, R)^* = (\pi(R^*), 1, \pi(L^*)), \quad (L, 0, R)^* = (L^*, 0, R^*).$$

Докажем индуктивно, что для непустого $T \in \mathcal{B}^i$ выполняется равенство $\lambda(T^*) = \rho(T) - i$. Если $T \simeq (\varnothing, i, \varnothing)$, то $T^* = T$, $\rho(T) = i$, $\lambda(T^*) = 0 = \rho(T) - i$. При доказательстве шага индукции (для $|T| \geqslant 2$) нужно отдельно рассматривать случаи, когда $L = \varnothing$, $R = \varnothing$ и $L, R \neq \varnothing$. Рассмотрим лишь последний случай: если $T \simeq (L, 1, R)$, то $\rho(T) = \rho(L) + \rho(R) + 1$ и

$$\lambda(T^*) = \lambda(R^*) + \lambda(L^*) + 1 = (\rho(R) - 1) + \rho(L) + 1 = \rho(T) - 1;$$

если $T \simeq (L, 0, R)$, то $\rho(T) = \rho(L) + \rho(R)$ и

$$\lambda(T^*) = \lambda(L^*) + \lambda(R^*) + 1 = (\rho(L) - 1) + \rho(R) + 1 = \rho(T).$$

Равенство $\lambda(T^*) = \rho(T) - 1$ для $T \in \mathcal{B}^1$ показывает, что деревья с k узлами нечетного ранга биективно соответствуют деревьям с $k-1$ левыми сыновьями. Ссылка на предложение 13 завершает доказательство.

Высотой узла a в упорядоченном дереве называется длина (число ребер) единственной простой цепи, соединяющей a с корнем.

Предложение 20. *Число деревьев из $\mathcal{T}_{n+1}$ с k узлами нечетной [четной] высоты равно $R(n, k)$.*

Доказательство. Утверждение следует из предложения 19, так как при биекции $\mathcal{B}_n \to \mathcal{T}_{n+1}$ из доказательства предложения 6 узлам нечетного ранга в прообразе отвечают узлы нечетной высоты в образе. Нечетную высоту на четную позволяет заменить симметрия чисел Рюньона.

Доказанное утверждение, видимо, известно: в [7, с. 368] говорится о «возможности установления взаимно-однозначного соответствия» между аналогами деревьев из предложений 17 и 20.

Решеточным многоугольником (или параллелограммным поли(о)мино, см. [10]) назовем пару (π, π') стандартных путей в целочисленной решетке (см. предложение 4), имеющих общее начало $(0,0)$ и общий конец и не имеющих других общих точек; π – «нижний», начинающийся с шага $(1,0)$, а π' – «верхний», начинающийся с шага $(0,1)$, пути. Абсцисса конечной точки называется *шириной* многоугольника. Очевидно, периметр такого многоугольника четный и равен удвоенной длине каждого пути. Следующий классический результат принадлежит Пойа [11]; доказательство адаптировано из [10]. Обозначим через PG_n множество всех решеточных многоугольников с периметром $2n$.

Предложение 21. (1) $|PG_{n+1}| = C_n$.
(2) *Число многоугольников из* PG_{n+1} *ширины* k *равно* $R(n,k)$.

Доказательство. Пусть

$$P = (\pi, \pi') \in PG_{n+1}, \quad \pi = (\pi_0, \pi_1, \ldots, \pi_n), \quad \pi' = (\pi'_0, \pi'_1, \ldots, \pi'_n).$$

Шаги π_0, π'_0, π_n, π'_n очевидны и не зависят от P. Остальные шаги составляют $n-1$ пар (π_i, π'_i), $i \in \langle n-1 \rangle$. Сопоставим паре (π_i, π'_i) букву ℓ, если $\pi_i = \pi'_i = (0,1)$ (пути «идут вверх»); букву r, если $\pi_i = \pi'_i = (1,0)$ («вправо»); букву x, если $\pi_i = (1,0)$, $\pi'_i = (0,1)$ (пути «расходятся»); букву $\overline{x}$, если $\pi_i = (0,1)$, $\pi'_i = (1,0)$ (пути «сходятся»). Легко видеть, что это – биекция PG_{n+1} на $\mathcal{M}_{n-1}$: каждое слово Моцкина – инструкция построения многоугольника. Отсюда следует (1). Наконец, ширина k – это суммарное число вхождений x и r, так как появление этих и только этих букв продвигает конструкцию на шаг вправо; поэтому (2) следует из предложения 12.

Следующие два утверждения Г. П. Егорычев ([12, с. 57, 58]) доказывает своим методом коэффициентов в рамках теории комплексных производящих функций.

Предложение 22. *Рассмотрим пары* (A, B) k*-подмножеств из* $\langle n \rangle$*, записанных по возрастанию элементов* $A = \{i_1, \ldots, i_k\}$, $B = \{j_1, \ldots, j_k\}$ *и обладающих свойством* $i_m \leqslant j_m$ *для всех* $m \in \langle k \rangle$*. Число таких пар равно* $R(n+1, k+1)$.

Доказательство. Установим биекцию перечисляемого множества на множество многоугольников из PG_{n+2} ширины $k+1$. Тогда утверждение будет следовать из предложения 21. Такой многоугольник составляется из $k+1$ столбиков ширины 1; верхняя ордината $(k+1)$-го столбика равна $n-k+1$.

Паре (A, B) из предложения сопоставим многоугольник, для которого числа $0, i_1 - 1, i_2 - 2, \ldots, i_k - k$ суть нижние ординаты столбиков, а числа $j_1, j_2 - 1, \ldots, j_k - k + 1, n - k + 1$ суть верхние ординаты столбиков. Легко видеть, что это биекция.

Как и для чисел $C(n,k)$, имея очевидное равенство $C_n = \sum_{k=1}^{n} R(n,k)$, получим обратное преобразование.

Предложение 23.

$$R(n+1, k+1) = \sum_{m=\lfloor k/2 \rfloor}^{n-k} \binom{n}{k+m} \binom{k+m}{k-m} C_m.$$

Доказательство. Согласно предложению 12, $R(n+1, k+1)$ равно числу слов $w \in \mathcal{M}_n$ с суммарным числом k вхождений x и ℓ. Если $d_x(w) = d_{\overline{x}}(w) = m$, то $d_l(w) = k - m$. Поэтому, чтобы сформировать такое слово, нужно из n мест выбрать $k + m$ мест для букв x, $\overline{x}$ и ℓ; из этих $k + m$ мест выбрать $k - m$ мест для ℓ, а на других $2m$ местах разместить слово из $\mathcal{D}_m$; наконец, просуммировать по m, что и приводит к требуемой формуле.

Заметим, что в данной работе намеренно опущены (чтобы не увеличивать объем) многие элементарные модели чисел Каталана (триангуляции выпуклых многоугольников, разложения чисел, ступенчатые коды, конфигурации точек на окружности и пр.). Кроме того, можно рассматривать другие комбинаторные функции, подчиненные числам Каталана, например, функцию

$$g(n,k) = \frac{1}{n} 2^{n-2k+1} \binom{n}{k} \binom{n-k}{k-1}$$

(см. [13]). Возможно, этому следует посвятить отдельную статью.

1. Каллеман Ж. Полугруппы и комбинаторные приложения. М.: Мир, 1985.

2. Коряков И. О. Сильно префиксные коды и обобщенная формула Рэйни // Изв. Урал. гос. ун-та. 2010. № 74. (Математика. Механика. Информатика; Вып. 12). С. 57–66.

3. Raney G. N. Functional composition patterns and power series reversion // Trans. Amer. Math. Soc. 1960. Vol. 94, № 5. P. 441–451.

4. Кнут Д. Искусство программирования для ЭВМ. Т. 1. Основные алгоритмы. М.: Мир, 1976.

5. Cori R., Vauquelin B. Planar maps and well labeled trees // Canad. J. Math. 1981. Vol. 23, № 5. P. 1023–1042.

6. Гарднер М. Путешествие во времени. М.: Мир, 1990.

7. Гульден Я., Джексон Д. Перечислительная комбинаторика. М.: Наука, 1990.

8. Феллер В. Введение в теорию вероятностей и ее приложения. М.: Мир, 1984. Т. 1.

9. Риордан Дж. Комбинаторные тождества. М.: Наука, 1982.

10. Делест М.-П., Вьенно Ж. Алгебраические языки и перечисление полимино // Кибернет. сб. Вып. 26. М.: Мир, 1989. С. 113–156.

11. Pólya G. On the number of certain lattice polygons // J. Comb. Theory. 1969. Vol. 6. P. 102–105.

12. Егорычев Г. П. Интегральное представление и вычисление комбинаторных сумм. Новосибирск: Наука, 1977.

13. Коряков И. О. Конечные префиксные коды и финитарные деревья. (В печати).

Статья поступила 06.12.2007